Systematic sampling locations for detecting an area of elevated values (hot spot)

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

| SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Detect the presence of a hot spot that has a specified size and shape
Type of Sampling Design	Hot spot
Sample Placement (Location) in the Field	Systematic (Hot Spot) with a random start location
Formula for calculating number of sampling locations	Singer and Wickman algorithm
Calculated total number of samples	10
Number of samples on map [a]	10
Number of selected sample areas [b]	31
Specified sampling area [c]	50817.00 ft^2
Grid pattern	Triangular
Size of grid / Area of grid cell [d]	77.707 feet / 5229.39 ft^2
Total cost of sampling [e]	$6000.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Size of grid / Area of grid cell gives the linear and square dimensions of the grid used to systematically place samples.
[e] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.
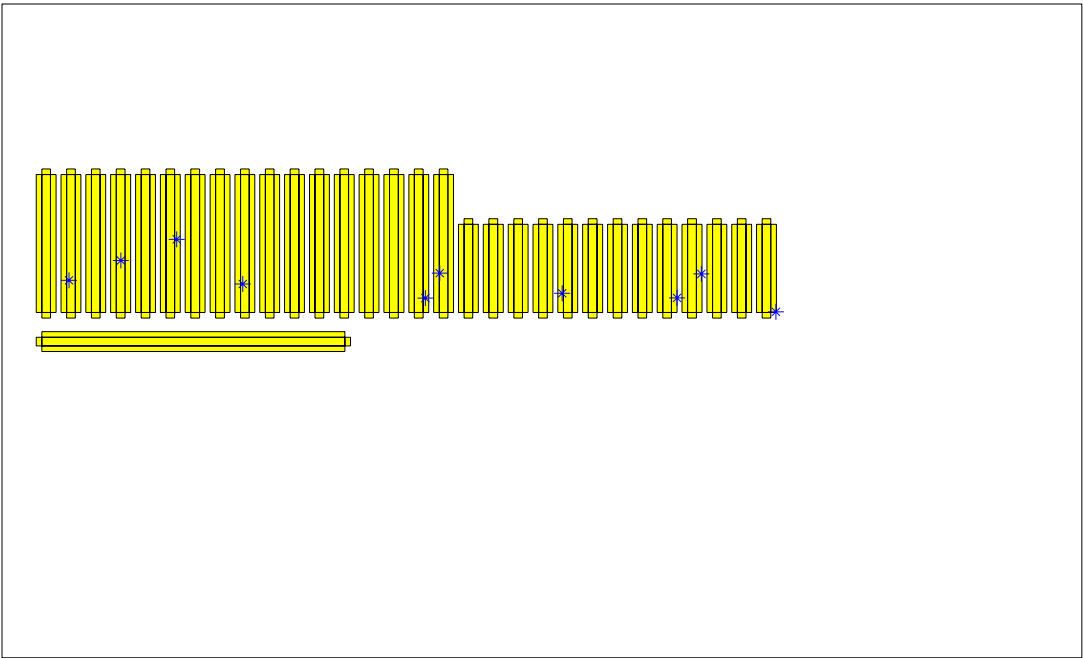
Area 1

X Coord	Y Coord	Label	Value	Type	Historical

Area 2

X Coord	Y Coord	Label	Value	Type	Historical
21.9563	25.8549		0	Hotspot	

Area 3

X Coord	Y Coord	Label	Value	Type	Historical

Area 4

X Coord	Y Coord	Label	Value	Type	Historical
63.6568	41.7998		0	Hotspot	

Area 5

X Coord	Y Coord	Label	Value	Type	Historical

Area 6

X Coord	Y Coord	Label	Value	Type	Historical
108.6185	58.7997		0	Hotspot	

Area 7

X Coord	Y Coord	Label	Value	Type	Historical

Area 8

X Coord	Y Coord	Label	Value	Type	Historical

Area 9

X Coord	Y Coord	Label	Value	Type	Historical
161.7947	22.8877		0	Hotspot	

Area 10

X Coord	Y Coord	Label	Value	Type	Historical

Area 11

X Coord	Y Coord	Label	Value	Type	Historical

Area 12

X Coord	Y Coord	Label	Value	Type	Historical

Area 13

X Coord	Y Coord	Label	Value	Type	Historical

Area 14

X Coord	Y Coord	Label	Value	Type	Historical

Area 15

X Coord	Y Coord	Label	Value	Type	Historical

Area 16

X Coord	Y Coord	Label	Value	Type	Historical
308.8000	11.5240		0	Hotspot	

Area 17

X Coord	Y Coord	Label	Value	Type	Historical
320.4580	31.6120		0	Hotspot	

Area 18

X Coord	Y Coord	Label	Value	Type	Historical

Area 19

X Coord	Y Coord	Label	Value	Type	Historical

Area 20

X Coord	Y Coord	Label	Value	Type	Historical

Area 21

X Coord	Y Coord	Label	Value	Type	Historical

Area 22

X Coord	Y Coord	Label	Value	Type	Historical
419.0233	15.3959		0	Hotspot	

Area 23

X Coord	Y Coord	Label	Value	Type	Historical

Area 24

X Coord	Y Coord	Label	Value	Type	Historical

Area 25

X Coord	Y Coord	Label	Value	Type	Historical

Area 26

X Coord	Y Coord	Label	Value	Type	Historical
511.3960	11.6992		0	Hotspot	

Area 27

X Coord	Y Coord	Label	Value	Type	Historical
531.0174	31.0138		0	Hotspot	

Area 28

X Coord	Y Coord	Label	Value	Type	Historical

Area 29

X Coord	Y Coord	Label	Value	Type	Historical

Area 30					
X Coord	Y Coord	Label	Value	Type	Historical
591.0915	0.5393		0	Hotspot	

Area 31					
X Coord	Y Coord	Label	Value	Type	Historical

Primary Sampling Objective

The primary purpose of sampling at this site is to detect "hot spots" (local areas of elevated concentration) of a given size and shape with a specified probability, $1-\beta$.

Selected Sampling Approach

This sampling approach requires systematic grid sampling with a random start. If a systematic grid is not used, the probability of detecting a hot spot of a given size and shape will be different than desired or calculated.

Number of Total Samples: Calculation Equation and Inputs

The algorithm used to calculate the grid size (and hence, the number of samples) is based on work by Singer for locating geologic deposits [see Singer (1972, 1975) and PNNL-13450 for details]. Inputs to the algorithm include the size, shape, and orientation of a hot spot of interest, an acceptable probability of not finding a hot spot, the desired type of sampling grid, and the sampling budget. For this design, the grid size was calculated based on a given hot spot size and other parameters.

The inputs to the algorithm that result in the grid size are:

| Parameter	Description	Value
Inputs		
$1-\beta$	Probability of detection	95%
Grid Type	Grid pattern (Square, Triangular or Rectangular)	Triangular
Hot Spot Shape	Hot spot height to width ratio	1
Hot Spot Size	Length of hot spot semi-major axis	40.22 feet
Hot Spot Area [a]	Area of hot spot (Length^2 * Shape * π)	5082 ft^2
Angle	Angle of orientation between hot spot and grid	Random
Sampling Area	Total area to sample	50817.00 ft^2
Outputs		
Grid Size	Length of side of grid	77.707 feet
Grid Area	Area covered by one grid cell	5229.39 ft^2
Samples [b]	Optimum number of samples	9.71757

[a] Length of semi-major axis is used by algorithm. Hot spot area is provided for informational purposes.
[b] The optimum number of samples is calculated by dividing the sampling area by the grid area.

The following graph shows the relationship between number of samples and the probability of finding the hot spot. The dashed blue line shows the actual number of samples for this design (which may differ from the optimum number of samples because of edge effects).

Hotspot Sampling of 50817 Feet^2
40.22 foot radius round hotspot
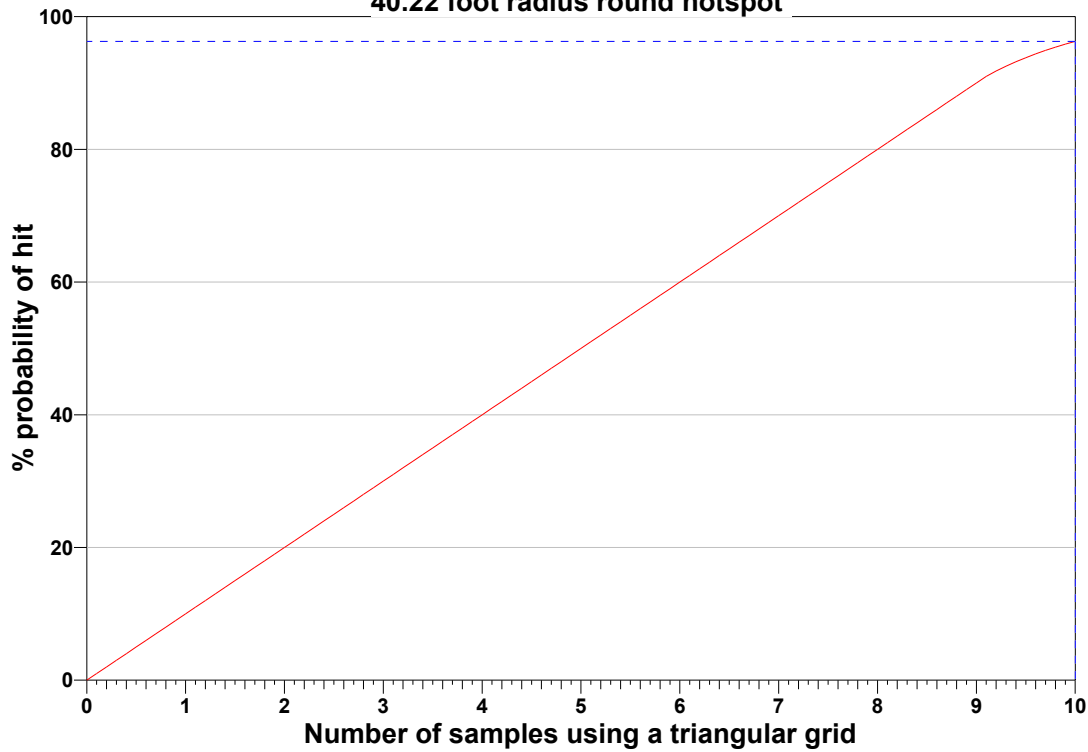
Statistical Assumptions

The assumptions associated with the sample spacing algorithm are that:

1. the target hot spot (its projection onto the coordinate plane) is circular or elliptical,
2. samples are taken on a square, rectangular, or triangular grid,
3. a very small proportion of the area being studied will be sampled (the sample is much smaller than the hot spot of interest),
4. the level of contamination that classifies a hot spot is well defined, and
5. there are no misclassification errors (a hot spot is not mistakenly overlooked or an area is not mistakenly identified as a hot spot).

These assumptions cannot be validated through data collection. The size and shape of a hot spot of interest are well defined prior to determining the number of samples and the measured value that defines a hot spot is well above the detection limit for the analytical methods that will be used. Grid sampling will be carried out to the level achievable; topographic, vegetative, and other features that prevent sampling at the specified coordinates will be noted and their influence recognized.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying $1-\beta$, Shp and Size and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

		Number of Samples		
		Size=20.11	Size=40.22	Size=60.33
1-β=90	Shp=0.8	47	12	6
	Shp=0.9	41	11	5
	Shp=1	36	9	4
1-β=95	Shp=0.8	52	13	6
	Shp=0.9	45	12	5
	Shp=1	39	10	5

	Shp=0.8	63	16	7
1-β=100	**Shp=0.9**	55	14	7
	Shp=1	49	13	6

$1-\beta$ = Probability of Hit (%)
Shp = Hot Spot Shape (Height to Width Ratio)
Size = Hot Spot Size (Length of Semi-major Axis)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $6000.00, which averages out to a per sample cost of $600.00. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**10 Samples**
Field collection costs		$100.00	$1000.00
Analytical costs	$400.00	$400.00	$4000.00
Sum of Field & Analytical costs		**$500.00**	**$5000.00**
Fixed planning and validation costs			$1000.00
Total cost			**$6000.00**

Recommended Data Analysis Activities
Post data collection activities generally follow those outlined in EPA's Guidance for Data Quality Assessment (EPA, 2000). The data analysts will become familiar with the context of the problem and goals for data collection and assessment. The data will be verified and validated before being subjected to statistical or other analyses. Graphical and analytical tools will be used to verify to the extent possible the assumptions of any statistical analyses that are performed as well as to achieve a general understanding of the data. The data will be assessed to determine whether they are adequate in both quality and quantity to support the primary objective of sampling.

A map of the actual sample locations will be generated so that the sampling plan and the field implementation may be compared. Deviations from planned sample locations due to topographic, vegetative, or other features will be noted. Their impacts will be qualitatively assessed. If a hot spot is discovered, additional sampling may be performed to determine its size and shape, in which case, the initial assumptions of the sampling design may then be assessed and/or reconsidered.